



Year: 2013

Utility maximization with a given pricing measure when the utility is not necessarily concave

Reichlin, Christian

Abstract: We study the problem of maximizing expected utility from terminal wealth for a not necessarily concave utility function

$$U$$

and for a budget set given by one fixed pricing measure. We prove the existence and several fundamental properties of a maximizer. We analyze the (not necessarily concave) value function (indirect utility)

$$u(x, U)$$

. In particular, we show that the concave envelope of

$$u(x, U)$$

is the value function

$$u(x, U_c)$$

of the utility maximization problem for the concave envelope

$$U_c$$

of the utility function

$$U$$

. The two value functions are shown to coincide if the underlying probability space is atomless. This allows us to characterize the maximizers for several model classes explicitly

DOI: <https://doi.org/10.1007/s11579-013-0093-x>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-156463>

Journal Article

Published Version

Originally published at:

Reichlin, Christian (2013). Utility maximization with a given pricing measure when the utility is not necessarily concave. *Mathematics and Financial Economics*, 7(4):531-556.

DOI: <https://doi.org/10.1007/s11579-013-0093-x>

Utility maximization with a given pricing measure when the utility is not necessarily concave

Christian Reichlin

Received: 2 June 2012 / Accepted: 2 January 2013 / Published online: 7 February 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract We study the problem of maximizing expected utility from terminal wealth for a not necessarily concave utility function U and for a budget set given by one fixed pricing measure. We prove the existence and several fundamental properties of a maximizer. We analyze the (not necessarily concave) value function (indirect utility) $u(x, U)$. In particular, we show that the concave envelope of $u(x, U)$ is the value function $u(x, U_c)$ of the utility maximization problem for the concave envelope U_c of the utility function U . The two value functions are shown to coincide if the underlying probability space is atomless. This allows us to characterize the maximizers for several model classes explicitly.

Keywords Portfolio selection · Non-concave utility · Asymptotic elasticity · Non-convex optimization · Behavioural finance

JEL Classifications G11

1 Introduction

For an increasing and upper-semicontinuous function U on \mathbb{R}_+ satisfying a mild growth condition, we study the problem

$$u(x, U) := \sup\{E[U(f)] \mid f \in C(x)\}, \quad (1.1)$$

where $C(x) := \{f \in L_+^0 \mid E_Q[f] \leq x\}$ for a (pricing) measure $Q \approx P$. Problem (1.1) arises naturally when one studies utility maximization. The function U can be seen as a (not necessarily concave) utility function describing the preferences of an agent in a financial

C. Reichlin (✉)
Department of Mathematics, ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland
e-mail: christian.reichlin@math.ethz.ch

C. Reichlin
Department of Banking and Finance, University of Zurich, Plattenstrasse 32, 8006 Zurich, Switzerland

market who is dynamically trading in the underlying discounted assets S . The random variable f can be seen as the final position $x + \int_0^T \vartheta dS$ resulting from a strategy ϑ with initial wealth x . In complete markets where there is a unique equivalent martingale measure Q for S , the elements in $C(x)$ are those non-negative positions that can be associated to a trading strategy starting from an initial endowment less than or equal to x . Then (1.1) is an abstract version of the utility maximization problem of choosing a trading strategy such that the terminal value f^* of the corresponding wealth process maximizes $E[U(f)]$ over all final outcomes of competing strategies. The idea for this reduction from a dynamic to a static problem can be traced back to Pliska [32]. The key advantage of the static formulation is that it describes in a simple and transparent way the fundamental economic problem. Moreover, it allows us to consider continuous- and discrete-time models in a unified way.

In the classical case where U is concave (which means that the agent is risk-averse) and smooth (e.g. in C^1), problem (1.1) and its solution are well known and we do not try to discuss it here; see Biagini [11] for a detailed survey. However, there is considerable empirical evidence (see for instance Tversky and Kahneman [40] and references therein) that agents tend to switch between risk-averse and risk-seeking behaviour, depending on the context. This leads to a partially convex (or convex–concave) utility function. Not necessarily concave utility functions also appear in the context of manager compensation [18], portfolio delegation [31], and strategic interaction among money managers [5].

In the case that U is not concave, problem (1.1) is more involved. There is a broad class of models in which the non-concave problem has been studied by reducing it to the classical concave case; see for instance Aumann and Perles [2], Carpenter [18], Berkelaar et al. [8], Larsen [31], Carassus and Pham [15], Rieger [35], Basak and Makarov [5] and Bichuch and Sturm [13]. At the other end of the scale, there are results on the existence of a solution in a number of (incomplete) discrete-time settings where one does not necessarily have a fixed pricing density, but the structure of the setup allows one to optimize directly over the set of strategies; see Benartzi and Thaler [7], Bernard and Ghossoub [10], He and Zhou [23] and Carassus and Rásonyi [16]. These two approaches provide fundamental results for a particular type of setting, but their economic implications are surprisingly varying.¹

The goal of this article is therefore to analyze (1.1) in a unified setting with one pricing density. From an economic point of view, this allows us to understand the behavioural predictions of non-concave parts in the utility function independently of the model-specific assumptions. In this way, we can explain and generalize in a systematic way the results obtained previously by other authors via ad hoc methods in specific frameworks. We also study an example in continuous time with jumps in the price process where the problem cannot be reduced to a concave one. We therefore not only unify but also extend the previous results in the literature. In contrast to the existing literature, we also present a detailed analysis of the value $u(x, U)$ as a function of the initial endowment x . This is a comparative static analysis to investigate the sensitivity of the optimal value with respect to the key exogenous variable “initial endowment”.

We present three main results. We first show necessary and sufficient conditions for the existence of a maximizer for $u(x, U)$, by using *Fatou’s lemma in several dimensions* due to Balder [3]; we also describe several fundamental properties of the maximizers. In particular, this requires the notion of asymptotic elasticity adapted to not necessarily concave utility

¹ One example is the relation between the optimal final position and the pricing density. While they are anti-comonotonic in the typical complete models in continuous time (Theorem B.1 in Jin and Zhou [27]), they are not necessarily anti-comonotonic in complete models in discrete time. Non-concave parts in the utility function might thus explain the pricing kernel puzzle in discrete time, but cannot do so in continuous time (see Sect. 6 in Hens and Reichlin [24] for a detailed discussion).

functions. We then study the value function $u(x, U)$. This is again an increasing and not necessarily concave function, and its concave envelope is shown to coincide with $u(x, U_c)$, which explains the general relation between the problem $u(x, U)$ and the concavified problem $u(x, U_c)$. In the third part, we specialize the setup to the case that the underlying probability space is atomless. Using rearrangement techniques, we show that $u(x, U)$ and $u(x, U_c)$ then coincide and we give a characterization of the maximizer for $u(x, U)$. We use our results to discuss explicit examples exhibiting a number of new and unexpected phenomena.

Besides the articles already mentioned above, there is another important branch of the literature that deals with non-concave problems. For more general preferences than expected utility, Jin and Zhou [27], Carlier and Dana [17] and He and Zhou [22] have developed (under the assumption that the distribution of dQ/dP is continuous) an approach via quantiles to obtain the existence (and in some cases also the structure) of a maximizer.

The paper is structured as follows. We start in Sect. 2 with a precise definition of the (not necessarily concave) utility function and its concave envelope. We then formulate the optimization problem and give an illustrative example which provides some basic intuition. In Sect. 3, we prove the existence and several properties of a maximizer. The value function is analyzed in Sect. 4. Section 5 contains a detailed analysis of the case that the underlying probability space is atomless. Finally, the appendices contain a number of well-known results from convex analysis and non-smooth utility maximization which are used in the body of the text.

2 Problem formulation and intuition

This section introduces the utility function, describes the framework and formulates the optimization problem we are interested in. For a probability space (Ω, \mathcal{F}, P) , let $L^0(\Omega, \mathcal{F}, P)$ (and $L^1(\Omega, \mathcal{F}, P)$) be the space of (equivalence classes of) \mathcal{F} -measurable (and integrable) random variables. The space $L^0_+(\Omega, \mathcal{F}, P)$ (and $L^1_+(\Omega, \mathcal{F}, P)$) consists of all non-negative elements of $L^0(\Omega, \mathcal{F}, P)$ (and $L^1(\Omega, \mathcal{F}, P)$). Sometimes, we drop the dependence on the probability space if it is clear from the context. For a random variable $f \in L^0(\Omega, \mathcal{F}, P)$, we use $f^\pm = \max(\pm f, 0)$ to denote the positive and negative parts of f .

Definition 2.1 A *utility* is a function $U : (0, \infty) \rightarrow \mathbb{R}$ with $U(\infty) > 0$, which is non-constant, increasing, upper-semicontinuous and satisfies the growth condition

$$\lim_{x \rightarrow \infty} \frac{U(x)}{x} = 0. \quad (2.1)$$

We only consider utility functions defined on the positive axis. To avoid any ambiguity, we set $U(x) = -\infty$ for $x < 0$ and define $U(0) := \lim_{x \searrow 0} U(x)$ and $U(\infty) := \lim_{x \nearrow \infty} U(x)$. Note that we do not assume that U is concave, continuous or strictly increasing. In particular, this also allows us to analyze goal-reaching problems initiated by Kulldorff [30]. In the concave case, condition (2.1) is equivalent to the Inada condition at ∞ that $U'(\infty) = 0$. The assumption $U(\infty) > 0$ is technical but completely harmless, because adding a constant to U does not change the preferences described by U . Moreover, note that the condition (2.1) and the assumption $U(\infty) > 0$ imply that there is always a concave function $g : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ satisfying $g \geq U$.

Definition 2.2 The *concave envelope* U_c of U is the smallest concave function $U_c : \mathbb{R} \rightarrow \mathbb{R} \cup \{-\infty\}$ such that $U_c(x) \geq U(x)$ holds for all $x \in \mathbb{R}$.

In order to formulate the optimization problem, we fix a probability space (Ω, \mathcal{F}, P) and consider a probability measure Q equivalent to P with density $\varphi := dQ/dP$. We refer to Q as *pricing measure* and to φ as *pricing density* (or *pricing kernel*). The set $C(x)$ is defined by

$$C(x) := \{f \in L_+^0 \mid E_Q[f] \leq x\}.$$

In this paper, we study the problem

$$u(x, U) := \sup\{E[U(f)] \mid f \in C(x)\}, \quad (2.2)$$

where we define $E[U(f)] := -\infty$ if $U^-(f) \notin L^1$. An element $f \in C(x)$ is *optimal* if $E[U(f)] = u(x, U)$. By a *maximizer* for $u(x, U)$, we mean an optimal element for the problem (2.2). The function $u(\cdot, U)$ is called the *value function* of the problem (2.2).

Remark 2.3 In behavioural finance, payoffs are often evaluated with respect to a (possibly stochastic) reference point R , which means that the agent evaluates $U(f - R)$ rather than $U(f)$. If the reference point is deterministic, this can be embedded in our analysis by defining a new utility $U_1(x) := U(x - R)$ and slightly modifying the arguments to account for the new domain of the utility. If the reference point is stochastic, one can solve the problem by maximizing first $E[U(g)]$ subject to the constraint $E_Q[g] \leq x - E_Q[R]$ to get a maximizer g^* and then choose $f^* := g^* + R$. In the setting of a complete financial market, in particular, this has a clear economic interpretation: the agent uses the amount $E_Q[R]$ to replicate the stochastic reference point R and invests the remaining part $x - E_Q[R]$ as if he had a deterministic reference point $E_Q[R]$. In particular, having a reference point in a complete market does not pose any significant difficulty.

2.1 Examples and intuition

We present here three representative models to illustrate how the pricing density and the underlying probability space look in explicit settings.

Example 2.4 [Black–Scholes model] We fix some time horizon $T \in (0, \infty)$, a probability space (Ω, \mathcal{F}, P) on which there is a standard Brownian motion $W = (W_t)_{t \geq 0}$ and a (discounted) market consisting of a savings account $B \equiv 1$, assumed to be constant, and one stock S described by

$$dS_t = S_t(\mu_t dt + \sigma_t dW_t), \quad S_0 = s_0 > 0, \quad \sigma > 0,$$

in the augmented filtration generated by W . The drift μ and the volatility σ are assumed to be adapted and the market price of risk $\lambda := \mu/\sigma$ is assumed to satisfy $\int_0^T \lambda_s^2 ds < \infty$ P -a.s. The unique martingale measure, if it exists, is defined by the pricing density $\varphi := dQ/dP = \exp\left(-\int_0^T \lambda_s dW_s - \frac{1}{2} \int_0^T \lambda_s^2 ds\right)$.

Example 2.5 [Geometric Poisson process] We consider some time horizon $T \in (0, \infty)$ and a probability space (Ω, \mathcal{F}, P) on which there is a Poisson process $N = (N_t)_{t \geq 0}$ with constant intensity $\lambda > 0$. Let $(\mathcal{F}_t)_{0 \leq t \leq T}$ be the augmented filtration generated by N . We consider a (discounted) market consisting of a savings account $B \equiv 1$ and one risky stock S described by

$$dS_t = \alpha S_t dt + \sigma S_t d\tilde{N}_t, \quad S_0 = s_0 > 0, \quad \sigma > -1, \quad \sigma \neq 0, \quad \alpha/\sigma < \lambda,$$

where $\tilde{N}_t := N_t - \lambda t$ is the compensated Poisson process. The unique martingale measure is defined by $\varphi := dQ/dP = e^{(\lambda - \tilde{\lambda})T} (\frac{\tilde{\lambda}}{\lambda})^{N_T}$ for $\tilde{\lambda} := \lambda - \alpha/\sigma$.

Example 2.6 [Complete model in discrete time] The classical example in discrete time is the binomial model consisting of a savings account $B \equiv 1$, assumed to be constant, and one stock S described by

$$\frac{S_k^1}{S_{k-1}^1} = Y_k = \begin{cases} 1 + u & \text{with probability } p \\ 1 + d & \text{with probability } 1 - p \end{cases}$$

for $k = 1, \dots, T$ in the filtration generated by Y . For $u > 0 > d > -1$, this model is complete; the underlying probability space consists of finitely many atoms.

The utility functions studied in this paper are not necessarily concave. In order to provide some intuition for these non-classical utility functions, we start with a motivating example in which we compare Example 2.4 for $\mu = 0$ and $T = 1$ and Example 2.6 for $u = -d < 1$ and $T = 1$. In both cases, it follows that $\varphi = 1$; in the first case, the underlying probability space is atomless; in the second case, the underlying probability space consists of two atoms. Example 2.7 shows that the two optimization problems are fundamentally different even though the pricing densities are equal: the underlying probability space crucially affects the optimization problem $u(x, U)$. Example 2.7 is of course pathological, but it is nevertheless the simplest possible setting demonstrating both intuition and structure and abstracting them from technical complexities.

Example 2.7 We consider the utility function U defined by

$$U(x) := \begin{cases} 0, & 0 < x < 1, \\ 1, & x \geq 1. \end{cases}$$

The concave envelope is given by $U_c(x) = x$ on $(0, 1)$ and $U_c(x) = 1$ for $x \geq 1$; the goal is therefore to maximize $P[f \geq 1]$. Moreover, we set $\varphi \equiv 1$.

For the concave problem $u(x_0, U_c)$, the random variable $f \equiv x_0$ is optimal (as can be seen by Jensen's inequality, because $\varphi \equiv 1$) and gives $u(x_0, U_c) = U_c(x_0) = x_0$. The maximizer for $u(x_0, U_c)$ and the value function $u(x_0, U_c)$ are independent of the underlying probabilistic structure.

For $u(x_0, U)$ and $x_0 \in (0, 1)$, the payoff $f \equiv x_0$ is not optimal; it is better to have 1 with probability x_0 while having 0 with probability $1 - x_0$. If the probability space is atomless, one can generate any probability distribution; this allows us to choose $A \in \mathcal{F}$ with $P[A] = x_0$. The element $f = 1_A$ gives the expected utility $U(0)P[A^c] + U(1)P[A] = x_0$ which is equal to $u(x_0, U_c)$. This shows that f is also a maximizer for $u(x_0, U)$, since $U_c \geq U$.

To illustrate the other extreme case, consider $\Omega := \{\omega_1, \omega_2\}$, $\mathcal{F} := 2^\Omega$ and $P[\{\omega_1\}] = P[\{\omega_2\}] = \frac{1}{2}$. In this model, a set A with $P[A] = x \in (0, 1) \setminus \{\frac{1}{2}\}$ cannot be generated since the elements in $L_+^0(\Omega, \mathcal{F})$ take at most two values, each with probability $1/2$, and so $u(x, U) = 1_{\{x \in [1/2, 1)\}} 1/2 + 1_{\{x \geq 1\}}$.

The important insight from this example is that the underlying probability space crucially affects the optimization problem $u(x, U)$. In the atomic case, the value function $u(x, U)$ is not concave and (for some x) strictly below $u(x, U_c)$. In the atomless case, the value function $u(x, U)$ is concave and equal to $u(x, U_c)$. Note that this has nothing to do with whether or not the pricing density φ has a continuous distribution.

2.2 Some tools from convex analysis

This section summarizes several results which will be used throughout the paper. The proofs can be found in Appendix A.

Lemma 2.8 *The concave envelope U_c of U is finite, continuous on $(0, \infty)$ and satisfies the condition (2.1). The set $\{U < U_c\} := \{x \in \mathbb{R}_+ \mid U(x) < U_c(x)\}$ is open and its (countable) connected components are bounded (open) intervals. Moreover, U_c is locally affine on the set $\{U < U_c\}$, in the sense that it is affine on each of the above intervals.*

A key tool to study the relation between U and U_c is the convex conjugate of U defined by

$$J(y) := \sup_{x>0} \{U(x) - xy\}.$$

If U is not concave, then the concave envelope U_c is not strictly concave and the latter implies that the function J is no longer smooth; we therefore work with the subdifferential which is denoted by ∂J for the convex function J and by ∂U_c for the concave function U_c (for precise definitions, see Appendix A). The right- and left-hand derivatives of J are denoted by J'_+ and J'_- . The next lemma summarizes several properties which are proved in Appendix A.

Lemma 2.9 *The function J is convex, decreasing, finite on $(0, \infty)$ and satisfies $J(x) = \infty$ for $x < 0$. The utility U and its concave envelope U_c have the same convex conjugate. Moreover, it holds that*

$$U_c(x) - xy = J(y) \iff x \in -\partial J(y) \iff y \in \partial U_c(x). \quad (2.3)$$

In classical concave utility maximization, the asymptotic elasticity (AE) of the utility function is introduced in Kramkov and Schachermayer [29]. They show that an upper bound on $AE(U)$ is essentially necessary and sufficient to prove the existence of an optimizer. Deelstra et al. [19] showed that this condition may also be formulated for the asymptotic elasticity of the convex conjugate. This turns out to be useful for extending asymptotic elasticity to the case when U is not necessarily concave. We define

$$AE_0(J) := \limsup_{y \rightarrow 0} \sup_{q \in \partial J(y)} \frac{|q|y}{J(y)}.$$

The condition $AE_0(J) < \infty$ has useful implications, which are summarized in the next lemma. The equivalence goes back to Lemma 6.3 in Kramkov and Schachermayer [29]. In the present form, it is proved in Lemma 4.1 of Deelstra et al. [19]. The proof of the second part is deferred to Appendix A.

Lemma 2.10 *The asymptotic elasticity condition $AE_0(J) < \infty$ is equivalent to the existence of two constants $\gamma > 0$ and $y_0 > 0$ such that*

$$J(\mu y) \leq \mu^{-\gamma} J(y) \quad \text{for all } \mu \in (0, 1] \text{ and } y \in (0, y_0].$$

Moreover, if $AE_0(J) < \infty$ is satisfied, then there are two constants $x_0 > 0$ and $k > 0$ such that $0 \leq U_c(x) \leq kU(x)$ on (x_0, ∞) .

In most applications, the utility U is strictly concave and differentiable on (x, ∞) for x large enough. For such utilities, the required asymptotic elasticity condition could also be formulated in the classical way in terms of $AE(U)$ or $AE(U_c)$, and the classical interpretation of the condition via risk aversion then still applies.

3 Existence and properties of a maximizer

The goal of this section is to prove the existence of a maximizer for $u(x, U)$. We also discuss several properties such as uniqueness, first order condition for optimality, and co-monotonicity of the maximizer and the pricing density.

To exclude the trivial case, it is necessary to assume that $u(x, U) < \infty$ for some $x > 0$. Even in the case of concave utilities, this is not sufficient to guarantee the existence of a maximizer; see, for instance, Sect. 5 in Kramkov and Schachermayer [29] and Sect. 5 in Biagini and Guasoni [12]. We introduce a stronger assumption.

Assumption 3.1 $E[J(\lambda\varphi)] < \infty$ for all $\lambda > 0$.

In the case that $U = U_c$, sufficient conditions for this assumption are $u(x, U_c) < \infty$ for some $x > 0$ and $AE_0(J) < \infty$. The next lemma extends this to the case when U is not necessarily concave; it is proved in Appendix C.

Lemma 3.2 *Under the assumption $AE_0(J) < \infty$, Assumption 3.1 is equivalent to $u(x, U) < \infty$ for some $x > 0$.*

We now turn to the main result of this section.

Theorem 3.3 *Suppose that Assumption 3.1 is satisfied. For all $x_0 \in (0, \infty)$, there exists some $\hat{f} \in C(x_0)$ such that $u(x_0, U) = E[U(\hat{f})]$.*

We work directly with a maximizing sequence (f^n) and try to get a suitable subsequence. Due to the lack of concavity of U , Komlós-type arguments (see, for instance, Lemma 5.5 in Westray and Zheng [41]) do not work. For later applications, we prove a slightly more general statement.

Proposition 3.4 *Suppose that Assumption 3.1 is satisfied. Fix a sequence $x_n \rightarrow x > 0$ and consider a sequence (f^n) with $f^n \in C(x_n)$. There is some $\hat{f} \in C(x)$ such that $\limsup_{n \rightarrow \infty} E[U(f^n)] \leq E[U(\hat{f})]$.*

Proof (1) In the first part, we show that the family $(U^+(f^n))_{n \in \mathbb{N}}$ is uniformly integrable. If U is bounded from above, this is clear. Hence, we may assume $U(\infty) = \infty$. The sequence (x_n) is bounded by x_0 , say, and it follows that

$$\begin{aligned} E[U^+(f^n)1_{\{U^+(f^n) > \alpha\}}] &\leq E[(J^+(\lambda\varphi) + f^n\varphi\lambda)1_{\{U^+(f^n) > \alpha\}}] \\ &\leq E[J^+(\lambda\varphi)1_{\{U^+(f^n) > \alpha\}}] + \lambda x_0. \end{aligned}$$

Therefore, it is sufficient to show that for any $\lambda > 0$, we have

$$\lim_{\alpha \rightarrow \infty} \sup_n E[J^+(\lambda\varphi)1_{\{U^+(f^n) > \alpha\}}] = 0.$$

Since the single random variable $J^+(\lambda\varphi) \in L^1$ is trivially uniformly integrable, we only need to show that $\sup_n P[U^+(f^n) > \alpha] \rightarrow 0$ for $\alpha \rightarrow \infty$. For this, fix a sequence $\alpha_i \rightarrow \infty$ and let $\tilde{x} := \inf\{x > 0 \mid U(x) > 0\}$ denote the first point where the utility becomes positive. By definition of \tilde{x} and x_0 , we have $U^+(f^n) \leq U(\tilde{x} + f^n)$ and $f^n + \tilde{x} \in C(x_0 + \tilde{x})$. And since we have by Lemma 3.2 that $u(x, U_c) < \infty$ for all $x > 0$, we get

$$\sup_n P[U^+(f^n) > \alpha_i] \leq \sup_n \frac{E[U^+(f^n)]}{\alpha_i} \leq \frac{u(x_0 + \tilde{x}, U)}{\alpha_i} \rightarrow 0,$$

which completes the proof of the first part.

(2) By passing to a subsequence that realizes the \limsup , we can assume that the sequence $(E[U(f^n)])$ converges, and we denote the limit by γ_1 . In the same way, by passing to a further subsequence (again, relabelled as f^n) that realizes $\limsup E[\varphi f^n]$, we can also assume that $(E[\varphi f^n])$ converges to some $\gamma_2 \leq x$. By part (1), we know that the family $(U(f^n)^+)_n$ is uniformly integrable; the same is trivially true for $((\varphi f^n)^-)_n$. By Corollary 3.9 of Balder [3], it follows then that there exist $\hat{g}_1 \in L^1$ and $\hat{g}_2 \in L^1$ such that

$$E[\hat{g}_1] \geq \gamma_1 \quad \text{and} \quad E[\hat{g}_2] \leq \gamma_2, \quad (3.1)$$

and for a.e. $\omega \in \Omega$, there exists a subsequence $n_k(\omega)$ such that

$$\lim_{k \rightarrow \infty} U(f^{n_k(\omega)}(\omega)) = \hat{g}_1(\omega) \quad \text{and} \quad \lim_{k \rightarrow \infty} \varphi(\omega) f^{n_k(\omega)}(\omega) = \hat{g}_2(\omega). \quad (3.2)$$

Since $\varphi > 0$ P -a.s., we can define $\hat{f}(\omega) := \hat{g}_2(\omega)/\varphi(\omega)$. This gives $\hat{f}\varphi = \hat{g}_2$ and together with (3.1), it follows that $E[\varphi \hat{f}] = E[\hat{g}_2] \leq \gamma_2 \leq x$, which means that $\hat{f} \in C(x)$. Moreover, it follows from (3.2) that

$$\lim_{k \rightarrow \infty} f^{n_k(\omega)}(\omega) = \lim_{k \rightarrow \infty} \frac{\varphi(\omega) f^{n_k(\omega)}(\omega)}{\varphi(\omega)} = \frac{\hat{g}_2(\omega)}{\varphi(\omega)} = \hat{f}(\omega).$$

Together with upper-semicontinuity of U and (3.2) we obtain

$$U(\hat{f}(\omega)) \geq \lim_{k \rightarrow \infty} U(f^{n_k(\omega)}(\omega)) = \hat{g}_1(\omega).$$

Taking expectations and using (3.1) gives $E[U(\hat{f})] \geq E[\hat{g}_1] \geq \gamma_1$. We conclude that \hat{f} is a maximizer since $\hat{f} \in C(x)$. \square

The proof of Theorem 3.3 is now a direct application of Proposition 3.4.

Proof of Theorem 3.3 Consider a maximizing sequence $(f^n)_{n \in \mathbb{N}}$ in $C(x_0)$. Since U is increasing, we can assume that the constraint $E[\varphi f^n] = x_0$ is satisfied for each $n \in \mathbb{N}$. Proposition 3.4 gives some $\hat{f} \in C(x_0)$ such that $E[U(\hat{f})] \geq \lim_n E[U(f^n)] = u(x_0, U)$, which shows that \hat{f} is a maximizer. \square

Remark 3.5 Balder and Pistorius [4] provide an existence result for a multi-good consumption problem with a not necessarily concave utility function on \mathbb{R}_+^m . They impose on the utility function a growth condition that also involves the pricing density. By exploiting convex duality, our Assumption 3.1 relaxes their conditions and highlights that the classical assumptions via asymptotic elasticity are sufficient for the existence of a maximizer also in the case with one pricing measure and not necessarily concave utilities.

Having clarified the existence of a solution to $u(x, U)$, it is natural to ask about uniqueness. Here the answer is negative: *A maximizer for $u(x, U)$ is not necessarily unique.* Similarly to the case of concave utilities which are not strictly concave, we can manipulate the solution on those parts of $\text{dom}(U)$ where U_c is locally affine. This is illustrated in the next example.

Example 3.6 Take a sufficiently rich model with $\varphi \equiv 1$ and consider a utility function satisfying $\{U < U_c\} = (a, b) \cup (b, c)$ for which U_c is affine on (a, c) . Jensen's inequality shows that $f \equiv b$ is a maximizer for $u(b, U)$. But on the other hand, for a set $A \in \mathcal{F}$

satisfying $P[A] = (c - b)/(c - a)$ (which exists by the richness of the model), it follows that $a1_A + c1_{A^c}$ is in $C(b)$ and satisfies $E[U(a1_A + c1_{A^c})] = U(b)$, which means that $f' = a1_A + c1_{A^c}$ is also a maximizer. Lemma 5.9 below describes model classes where the solution for $u(x, U)$ is unique.

In the sequel, our goal is to describe the properties of maximizers in more detail. We start with the first order condition for optimality.

Lemma 3.7 *Let U be continuously differentiable with $\{U < U_c\} = \bigcup_{i=1}^n (a_i, b_i)$ for some fixed n and let f^* be a maximizer for $u(x, U)$. Then there is $\lambda > 0$ such that f^* satisfies $U'(f^*) = \lambda\varphi$ on $\{f^* > 0\}$.*

Proof Fix $\epsilon > 0$, define $A_\epsilon := \{f^* > \epsilon\}$ and consider

$$\max E[U(f)1_{A_\epsilon}] \quad \text{subject to } f \in L^0_+, E[f\varphi 1_{A_\epsilon}] \leq E[f^*\varphi 1_{A_\epsilon}]. \quad (3.3)$$

If there is some element \hat{f} in (3.3) with $E[U(\hat{f})1_{A_\epsilon}] > E[U(f^*)1_{A_\epsilon}]$, the candidate $f' := \hat{f}1_{A_\epsilon} + f^*1_{A_\epsilon^c}$ is feasible for the problem $u(x, U)$ and satisfies $E[U(f')] > E[U(f^*)]$ which contradicts the optimality of f^* . Hence f^* also solves (3.3). Now fix some $f \in L^\infty$ and define $f_\lambda := f^* + \lambda(f - c)1_{A_\epsilon}$ for $c = E[\varphi f 1_{A_\epsilon}]/E[\varphi 1_{A_\epsilon}]$. First, note that

$$E[\varphi f_\lambda 1_{A_\epsilon}] = E[\varphi f^* 1_{A_\epsilon}] + \lambda(E[\varphi f 1_{A_\epsilon}] - cE[\varphi 1_{A_\epsilon}]) = E[\varphi f^* 1_{A_\epsilon}]$$

holds for every λ . Moreover, $f \in L^\infty$ implies $f_\lambda \geq 0$ on A_ϵ for λ small. Hence f_λ is a feasible candidate for the problem (3.3) and this yields

$$0 \geq \limsup_{\lambda \rightarrow 0} \frac{E[(U(f_\lambda) - U(f^*))1_{A_\epsilon}]}{\lambda}.$$

Since U is continuously differentiable and concave on (b_n, ∞) , U' is bounded on (ϵ, ∞) , and using the mean value theorem, we find that $(U(f_\lambda) - U(f^*))/\lambda$ is bounded by a constant on A_ϵ . Hence we may interchange limit and expectation to obtain $0 \geq E[U'(f^*)(f - c)1_{A_\epsilon}]$. Replacing f by $-f$ shows that the expectation must vanish. Using the notation $\gamma := E[U'(f^*)1_{A_\epsilon}]/E[\varphi 1_{A_\epsilon}]$, we thus see that $E[(U'(f^*) - \varphi\gamma)f 1_{A_\epsilon}] = 0$ holds for all $f \in L^\infty$. This implies $U'(f^*) = \varphi\gamma$ on A_ϵ . The same approach for $\tilde{\epsilon} \in (0, \epsilon)$ gives $U'(f^*) = \varphi\tilde{\gamma}$ on $A_{\tilde{\epsilon}}$ for some constant $\tilde{\gamma}$. Since $A_\epsilon \subset A_{\tilde{\epsilon}}$, we have that $U'(f^*) = \varphi\gamma = \varphi\tilde{\gamma}$ on A_ϵ and we infer $\gamma = \tilde{\gamma}$. This can be done for any $\tilde{\epsilon} > 0$ and we obtain $U'(f^*) = \varphi\gamma$ on $\bigcup_{\epsilon > 0} A_\epsilon = \{f^* > 0\}$, which proves the assertion. \square

In financial economics, there is a broad strand of literature (see for instance Jackwerth [25] and Beare [6]) that analyzes the pricing density in financial markets. From an equilibrium perspective, this boils down to an analysis of the relation between the pricing density dQ/dP and the optimal final position f^* of a single agent. In the classical case where U is concave, the marginal utility U' is decreasing and the first order condition $U'(f^*) = \lambda dQ/dP$ hence gives a decreasing relation between f^* and dQ/dP . This means that f^* and dQ/dP are *anti-comonotonic* (see Definition 4.82 in Föllmer and Schied [21]). Dybvig [20] considers a finite-dimensional setup and shows, by using rearrangement techniques, that the decreasing relation between f^* and dQ/dP also holds for not necessarily concave utility functions if all states $\omega \in \Omega$ have the same probability. The economic intuition for this result is that it is (cost-)efficient to have a higher payoff in those states where the pricing density is low. Jin and Zhou [27] and Carlier and Dana [17] show that the decreasing relation between the optimal final position and dQ/dP also holds for more general preferences if the pricing density has a continuous distribution. However, this result cannot be generalized to an arbitrary

combination of pricing density and underlying probability space as we illustrate in the next example.

Example 3.8 Take $\Omega = \{\omega_1, \omega_2\}$, $\mathcal{F} = 2^\Omega$, $P[\{\omega_1\}] = 2/3$, define the pricing measure by $Q[\{\omega_1\}] = 3/4$ and consider the utility function

$$U(x) := \begin{cases} (x-1)^{1/3}, & x \geq 1, \\ -(1-x)^{2/3}, & x < 1. \end{cases}$$

The problem $u(1, U)$ can be solved explicitly for this function U (see Theorem 3.1 in Bernard and Ghossoub [10]), and it turns out that f^* defined by $f^*(\omega_1) := 10/9$ and $f^*(\omega_2) := 2/3$ is the optimal final position for $u(1, U)$. It follows that f^* and φ are not anti-comonotonic.

Remark 3.9 (1) By extending, for the case of expected utility, the results of Jin and Zhou [27] and Carlier and Dana [17], it is shown below in Corollary 5.6 that φ and the maximizer f^* for $u(x, U)$ are anti-comonotonic if the probability space is atomless.

(2) At first glance, Example 3.8 seems to contradict the main result of Rieger [34] who claims to generalize the above result in Dybvig [20] to a general pricing density. However, a closer inspection of the setup and the results in Rieger [34] shows that the optimization problem (Definitions 2.10 and 2.11 and Theorem 2.12 there) is not of the same classical form as our problem (2.2). Rieger [34] considers probability distributions where the pricing density φ and the final payoff f may have any joint distribution (in contrast to our setup, where only those joint distributions are allowed which can be actually supported by the underlying probability space), and the constraint is formulated in terms of the covariance between the payoff and the pricing density. In the context of Example 3.8, this allows one to choose a payoff f via the joint distribution

$$P[\varphi = 3/4, f = 10/9] = P[\varphi = 9/8, f = 10/9] = P[\varphi = 9/8, f = 2/3] = 1/3,$$

which is not possible in our framework.

One can think of the approach in Rieger [34] as allowing for *randomized* payoffs. In the above example, instead of a fixed payoff $f(\omega_2)$, one is allowed to choose a lottery with outcomes 10/9 and 2/3 with probability 0.5 each. For mathematical purposes, having randomized payoffs is very useful since it allows one to prove anti-comonotonicity of f^* and φ . But if the payoffs must be generated by trading in a specified financial market, the underlying probabilistic structure matters, and allowing randomized payoffs does not match up well with the given financial problem.

(3) Instead of the (primal) problem to maximize the expected utility, one can also look at a dual problem of minimizing over payoffs f the cost $E_Q[f]$ for a given value of $E[U(f)]$. Since the expected utility depends only on the distribution of f , this problem is closely related to finding the cheapest price to generate a given distribution. In Dybvig [20], this is called the *distributional price* of f or of its distribution, and a payoff is called *cost-efficient* if any other payoff that generates the same distribution costs at least as much. The ideas of Dybvig [20] have recently been revisited, formalized and extended in Bernard and Boyle [9]. One central result there is to show that a payoff is cost-efficient if the payoff and the pricing density are anti-comonotonic. Example 3.8 above shows that the converse direction is not true in general: A pricing density φ and a cost-efficient payoff are not necessarily anti-comonotonic.

4 The value function

In Sect. 3, we have considered $u(x, U)$ for a fixed $x > 0$. This section analyzes how the optimal value $u(x, U)$ depends on x . Not surprisingly, the value function $u(x, U)$ is again increasing, upper-semicontinuous and not necessarily concave (upper-semicontinuity is proved below in Proposition 4.2). So the value function itself has similar properties as the utility U and, as in the case of U , we are interested in the concave envelope of $u(x, U)$. The first important result is that the concave envelope of $u(x, U)$ is (essentially) given by $u(x, U_c)$. This gives a general relation between $u(x, U)$ and $u(x, U_c)$.

Theorem 4.1 *Take any utility function U with convex conjugate J . If Assumption 3.1 is satisfied, then the value function $u(x, U_c)$ and the concave envelope of $u(x, U)$ coincide on $(0, \infty)$ and we have*

$$j(\lambda) := E[J(\lambda\varphi)] = \sup_{x>0} \{u(x, U) - x\lambda\}, \quad \lambda > 0. \quad (4.1)$$

The relation (4.1) states that j is the conjugate of the value function $u(x, U)$. In the concave case, the corresponding result from the literature is more general since j there is defined as an infimum (over pricing measures) in a dual problem; see for instance Lemma 3.4 of Kramkov and Schachermayer [29]. However, a closer inspection of that proof there shows that concavity is not needed for our present case with a single pricing measure. More precisely, Kramkov and Schachermayer [29] (or Westray and Zheng [41] for the nonsmooth case) determine the conjugate of the value function in a setting with infinitely many pricing measures. Then concavity of the utility function U is (essentially) used to apply the minimax theorem in order to exchange the supremum over (bounded) payoffs and the infimum over pricing densities. In the present case with one single pricing density, this step is not necessary and one can proceed without concavity. For completeness, we carry out the details.

Proof of Theorem 4.1 (1) To show (4.1), we start with the claim that

$$\lim_{n \rightarrow \infty} \sup_{f \in C_n} E[U(f) - f\lambda\varphi] = \sup_{x>0} \{u(x, U) - x\lambda\}, \quad (4.2)$$

where $C_n := \{f \in L_+^0 \mid 0 \leq f \leq n\}$ is the ball of radius n in the positive orthant of L^0 . For “ \leq ”, note that the left-hand side of (4.2) is an increasing limit in n ; so we have to show that for each n and each $f \in C_n$,

$$E[U(f) - f\lambda\varphi] \leq \sup_{x>0} \{u(x, U) - x\lambda\}.$$

To do that, fix f and define $x^* := E[\varphi f]$. For $x^* = 0$, we get $E[f\varphi] = 0$, hence $f \equiv 0$, and so $E[U(f) - f\lambda\varphi] = U(0) \leq u(x, U) - x\lambda + x\lambda$ for any $x > 0$. This gives the above inequality, and so we consider the case where $x^* > 0$. By definition of x^* , we have $f \in C(x^*)$ and it follows that

$$E[U(f) - f\lambda\varphi] \leq u(x^*, U) - x^*\lambda \leq \sup_{x>0} \{u(x, U) - x\lambda\}.$$

This proves (4.2).

(2) To get (4.1) from (4.2), we now want to interchange supremum and expectation on the left-hand side of (4.2) and then let $n \rightarrow \infty$. For each n , a measurable selection argument (see Theorem 18.19 in Aliprantis and Border [1]) shows that we can choose a measurable selector $x^*(\omega) \in L_+^0$ such that

$$\sup_{0 \leq x \leq n} \{U(x) - x\lambda\varphi(\omega)\} = U(x^*(\omega)) - x^*(\omega)\lambda\varphi(\omega).$$

With $J_n(\lambda) := \sup_{0 \leq x \leq n} \{U(x) - x\lambda\} \geq U(n) - n\lambda$, it thus follows that we have indeed

$$\sup_{f \in C_n} E[U(f) - f\lambda\varphi] = E[J_n(\lambda\varphi)]. \quad (4.3)$$

But J_n is increasing in n and dominated by J , so in view of (4.2) and (4.3), we have to show for (4.1) that $\lim_{n \rightarrow \infty} E[J_n(\lambda\varphi)] \geq E[J(\lambda\varphi)]$. Because of $J_n^\pm \rightarrow J^\pm$ and Fatou's lemma, it is sufficient to show that $(J_n^-(\lambda\varphi))_{n \in \mathbb{N}}$ is uniformly integrable. Since J_n is increasing in n , J_n^- is decreasing in n and $J_1^-(\lambda\varphi) \leq |U(1)| + \lambda\varphi$ is an integrable upper bound for $J_n^-(\lambda\varphi)$, $n \in \mathbb{N}$.

(3) It remains to prove that $u(x, U_c)$ and the concave envelope of $u(x, U)$ coincide on $(0, \infty)$. Note first from Lemma 2.9 that U and U_c have the same convex conjugate J so that applying (4.1) for U and U_c implies that j is the conjugate of both $u(x, U)$ and $u(x, U_c)$. Thus also their biconjugates coincide. But applying part (iii) of Lemma A.1 to $u(x, U_c)$ gives that $u(x, U_c)$ is equal to its biconjugate. This shows that $u(x, U_c)$ is the biconjugate of $u(x, U)$. So applying part (iii) of Lemma A.1 now to $u(x, U)$ (which is upper-semicontinuous due to Proposition 4.2 below) finally gives that $u(x, U_c)$ and the concave envelope of $u(x, U)$ coincide on $(0, \infty)$. \square

Example 2.7 shows that the value function $u(x, U)$ is not continuous in general. If, however, the utility function is continuous, then also the value function $u(x, U)$ is continuous.

Proposition 4.2 *Let Assumption 3.1 be satisfied. Then the value function $u(x, U)$ is upper-semicontinuous. If U is, in addition, continuous, then the value function $u(x, U)$ is continuous on $(0, \infty)$.*

Proof For upper-semicontinuity, consider a sequence $x_n \searrow x \in (0, \infty)$ and, using Theorem 3.3, denote the maximizer for $u(x_n, U)$ by f^n . Proposition 3.4 gives $\hat{f} \in C(x)$ satisfying

$$\limsup_{n \rightarrow \infty} u(x_n, U) = \limsup_{n \rightarrow \infty} E[U(f^n)] \leq E[U(\hat{f})] \leq u(x, U),$$

which finishes the proof of upper-semicontinuity for $x \in (0, \infty)$. Moreover, it is known from the literature (Theorem 6 of Siorpaes [38]) that $u(x_n, U_c) \searrow U_c(0)$ for $x_n \searrow 0$ and this implies $u(x_n, U) \searrow U(0)$ for $x_n \searrow 0$ since $U_c(0) = U(0) \leq u(x, U) \leq u(x, U_c)$ for all $x > 0$.

It remains to show the lower-semicontinuity of $u(x, U)$ if U is continuous. Fix a sequence $x_n \nearrow x_0 \in (0, \infty)$ and a random variable $f_0 \in C(x_0)$ with $U^-(f_0) \in L^1$. Then we construct a sequence $(f_n)_n$ with $f_n \in C(x_n)$, $f_n \nearrow f_0$ and $U^-(f_n) \in L^1$ for all n . Dominated convergence for $(U^-(f_n))_n$ (with the upper bound $U^-(f_1)$) and Fatou's lemma for $(U^+(f_n))_n$ give

$$E[U(f_0)] \leq \liminf_{n \rightarrow \infty} E[U(f_n)] \leq \liminf_{n \rightarrow \infty} u(x_n, U).$$

The result follows by taking the supremum over $f_0 \in C(x)$. \square

5 Results for an atomless model

In this section, we specialize our setup to the case when the underlying probability space is atomless. This is equivalent to assuming that the probability space supports a continuous distribution (see Definition A.26 and Proposition A.27 in Föllmer and Schied [21] for a precise definition and equivalent formulations). One special example of this occurs when

we have a pricing density with a continuous distribution. This case often appears in the literature; see for instance Larsen [31]. It is shown there in the setting of the Black–Scholes model that $u(x, U_c) = u(x, U)$ and that the utility maximization problem $u(x, U)$ can be reduced to the concavified problem $u(x, U_c)$. However, the assumption that the distribution of φ is continuous is not very satisfactory from an economic point of view since it does not provide any structural explanation for the results. Also, from a mathematical point of view, the assumption is not very elegant since it might be tedious or even impossible to verify the continuity of the distribution. Carassus and Pham [15] and Bichuch and Sturm [13] use techniques from Malliavin calculus (and additional assumptions) to show the continuity of the distribution of φ in Example 2.4; we also present an example with jumps in the price process where the underlying probability space is atomless and the distribution of the pricing density is not continuous.

The main result of this section is that $u(x, U) = u(x, U_c)$ already holds if the underlying probability space is atomless. This unifies and generalizes the existing results in the literature, and the proof also provides an economically intuitive explanation. Note also that the assumption of having an atomless probability space is easily verifiable in applications.

Theorem 5.1 *Let (Ω, \mathcal{F}, P) be atomless. Then it holds that*

$$u(x, U) = u(x, U_c) \quad \text{for all } x > 0. \quad (5.1)$$

The problem $u(x, U)$ admits a maximizer if and only if the concavified problem $u(x, U_c)$ admits a maximizer. Every maximizer for the problem $u(x, U)$ also maximizes the concavified problem $u(x, U_c)$.

While every maximizer for $u(x, U)$ is also a maximizer for $u(x, U_c)$, the opposite conclusion is not true in general.

Example 5.2 *A maximizer for $u(x, U_c)$ is in general not a maximizer for $u(x, U)$. Fix $\varphi \equiv 1$ and $x \in \{U < U_c\}$. Doing nothing, i.e., $f \equiv x$, is optimal for $u(x, U_c)$, but is not optimal for $u(x, U)$. Lemma 5.7 below describes model classes where a solution for $u(x, U_c)$ is also a solution for $u(x, U)$.*

As we see below in Lemma 5.7, Theorem 5.1 is straightforward and (essentially) known in the case in which the distribution of φ is continuous. In general, it is therefore a natural idea to approximate possible mass points in that distribution by continuous distributions. Such an approach is possible, but it requires some new and involved convergence results as well as additional restrictive integrability conditions on φ (see Sect. 3.3 of Reichlin [33]). We use another approach based on rearrangement techniques that works without additional assumptions on φ . This approach is based on a slightly stronger result.

Proposition 5.3 *Suppose that (Ω, \mathcal{F}, P) is atomless. For any candidate $f \in C(x)$, there exists $f^* \in C(x)$ satisfying $\{f^* \in \{U < U_c\}\} = \emptyset$ and*

$$E[U(f^*)] = E[U_c(f^*)] = E[U_c(f)] \geq E[U(f)], \quad (5.2)$$

where the last inequality is strict if and only if $P[f \in \{U < U_c\}] > 0$.

This proposition is inspired by Rieger [35] who solved the case of a pricing density with a continuous distribution. In addition to giving a rigorous proof, our contribution compared to Rieger [35] is to prove the statement without any assumption on the distribution of φ . This is the crucial step in extending the existing results in the literature. Let us first explain the idea. Fix a utility U with $\{U < U_c\} = (a, b)$ and choose $x \in (a, b)$. We consider the random

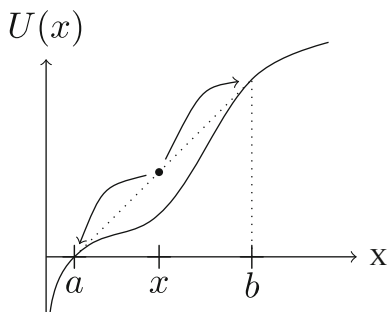


Fig. 1 concentrate all the weight in (a, b) on the corners a and b

variable $f \equiv x$. The idea is to use *rearrangement techniques* to construct $f^* \in C(x)$ with the same expected concavified utility and no probability weight in $\{U < U_c\}$. The method to achieve this is to concentrate all the weight on a and b (see Fig. 1), i.e., to choose f^* of the form $f^* = a1_A + b1_{A^c}$. The set $A \in \mathcal{F}$ has to be chosen in such a way that the agent can still afford the claim.

In the special case $\varphi \equiv 1$, a feasible choice for A is any set satisfying $P[A] = \frac{b-x}{b-a}$. For an arbitrary φ , the condition on A is a bit more involved; in particular, we have to put the “expensive” states ($\varphi(\omega)$ high) on a and the “cheap” states ($\varphi(\omega)$ low) on b . For a general utility U with $\{U < U_c\} = \bigcup_i (a_i, b_i)$ as in Lemma 2.8, every set $\{\omega \in \Omega \mid f \in (a_i, b_i)\}$ is rearranged separately in a similar way by concentrating the weight on a_i and b_i . For $\omega \in \Omega$ with $f(\omega) \in \{U = U_c\}$, the random variable $f^*(\omega) := f(\omega)$ is not changed.

Proof of Proposition 5.3 We split the proof into several steps; the plan is as follows. In the first part, we start with some preliminary remarks and define the sets needed for the construction of f^* . We then construct f^* and show that $f^* \in C(x)$ in the second part. (5.2) is shown in the third part. The final “if and only if” statement is a direct consequence of (5.2).

- (1) It is shown in Lemma 2.8 that the concave envelope U_c is locally affine on $\{U < U_c\}$ and that $\{U < U_c\} = \bigcup_i (a_i, b_i)$ for some a_i and b_i . We define

$$S_i := \{\omega \in \Omega \mid f(\omega) \in (a_i, b_i)\}$$

and $S := \bigcup_i S_i = \{f \in \{U < U_c\}\}$, which means that S contains all the states where f takes values in the non-concave part of U . For $\omega \in S_i$, the weight $\lambda(\omega)$ is defined by $\lambda(\omega) := (b_i - f(\omega))/(b_i - a_i)$, and since $f(\omega) \in (a_i, b_i)$ by definition of S_i , it holds that $\lambda(\omega) \in (0, 1)$. An elementary calculation shows that for $\omega \in S_i$, we have (by construction)

$$f(\omega) = \lambda(\omega)a_i + (1 - \lambda(\omega))b_i. \quad (5.3)$$

The idea is now to decompose every set S_i into two parts. Because of the atomless structure of the probability space, there exists (Föllmer and Schied [21, Lemma A.28]) a random variable \mathcal{U} with a uniform distribution on $(0, 1)$ and an increasing function q_φ on $(0, 1)$ such that $\varphi = q_\varphi(\mathcal{U})$ holds P -a.s. (q_φ is a *quantile* function of φ). Since $f_i(s) := P[S_i \cap \{\mathcal{U} < s\}]$ is a continuous function from $[0, 1]$ to $[0, P[S_i]]$, we find s_i such that $f_i(s_i) = E[(1 - \lambda)1_{S_i}]$. We now define the disjoint sets

$$S_{i1} := S_i \cap \{\mathcal{U} \geq s_i\} \quad \text{and} \quad S_{i2} := S_i \cap \{\mathcal{U} < s_i\}.$$

Note that $S_{i1} \cup S_{i2} = S_i$, which gives $P[S_{i1}] = P[S_i] - P[S_{i2}]$. Moreover, due to the definition of s_i , we can express the probabilities of S_{i2} and S_{i1} in terms of S_i and λ . More precisely, we have

$$P[S_{i2}] = f_i(s_i) = E[(1 - \lambda)1_{S_i}] \text{ and } P[S_{i1}] = E[\lambda 1_{S_i}]. \quad (5.4)$$

Finally, note that $\mathcal{U} < s_i$ on S_{i2} and $\mathcal{U} \geq s_i$ on S_{i1} by definition. Since $\varphi = q_\varphi(\mathcal{U})$ is an increasing function of \mathcal{U} , we have

$$\sup_{\omega \in S_{i2}} \varphi(\omega) \leq \inf_{\omega \in S_{i1}} \varphi(\omega). \quad (5.5)$$

- (2) The modified random variable f^* is defined by

$$f^*(\omega) := \begin{cases} f(\omega), & \omega \in S^c, \\ a_i, & \omega \in S_{i1}, \\ b_i, & \omega \in S_{i2}. \end{cases}$$

The measurability of f implies that $S_i = \{f \in (a_i, b_i)\} \in \mathcal{F}$, and the measurability of \mathcal{U} implies that $\{\mathcal{U} \geq s_i\} \in \mathcal{F}$; so also the intersection $S_{i1} = S_i \cap \{\mathcal{U} \geq s_i\}$ is in \mathcal{F} . Since $\{f^* \leq x\}$ can be written in terms of f and S_{i1} , we get measurability of f^* . We show below in step 4) that

$$a_i E[\varphi 1_{S_{i1}}] + b_i E[\varphi 1_{S_{i2}}] \leq E[\varphi(\lambda a_i + (1 - \lambda)b_i) 1_{S_i}], \quad (5.6)$$

which means that $E[\varphi f^* 1_{S_i}] \leq E[\varphi f 1_{S_i}]$ holds for every i . Since we have $f^* = f$ on $S^c = (\bigcup_i S_i)^c$, the assumption $f \in C(x)$ then gives $E[\varphi f^*] \leq E[\varphi f] \leq x$, which means that $f^* \in C(x)$.

- (3) For the first equality in (5.2) that $E[U(f^*)] = E[U_c(f^*)]$, note that $f^* = f$ on S^c by the definition of f^* . On $S = \bigcup_i S_i$, the definition of f^* gives $f^* \in \{a_i, b_i\}$ on S_i . By Lemma 2.8, the constants a_i and b_i are also in $\{U = U_c\}$; hence we arrive at $f^*(\omega) \in \{U = U_c\}$ for all $\omega \in \Omega$. The latter implies $E[U(f^*)] = E[U_c(f^*)]$. For the second equality in (5.2) that $E[U_c(f^*)] = E[U_c(f)]$, we first show that $E[U_c(f) 1_{S_i}] = E[U_c(f^*) 1_{S_i}]$ holds for every i . This needs four ingredients. In the first step, we apply (5.3) to rewrite f in terms of a_i , b_i and λ . Second, we use that the concave envelope U_c is affine on $[a_i, b_i]$ and the fact that $\lambda(\omega) \in (0, 1)$. We then apply (5.4) and finally, we rewrite the resulting convex combination in terms of f^* . Following these steps, we obtain

$$\begin{aligned} E[U_c(f) 1_{S_i}] &= E[U_c(\lambda a_i + (1 - \lambda)b_i) 1_{S_i}] \\ &= U_c(a_i)E[\lambda 1_{S_i}] + U_c(b_i)E[(1 - \lambda) 1_{S_i}] \\ &= U_c(a_i)P[S_{i1}] + U_c(b_i)P[S_{i2}] \\ &= E[U_c(f^*) 1_{S_i}] \end{aligned}$$

for every i . Recall that $S = \bigcup_i S_i$ holds by definition and that $f = f^*$ on the complement S^c by the definition of f^* . Monotone convergence and the above equality show that $U_c^\pm(f) \in L^1$ if and only if $U_c^\pm(f^*) \in L^1$ and that

$$E[U_c(f) 1_S] = \sum_i E[U_c(f) 1_{S_i}] = \sum_i E[U_c(f^*) 1_{S_i}] = E[U_c(f^*) 1_S]$$

if $U_c^\pm(f) \in L^1$. In all cases, we obtain $E[U_c(f)] = E[U_c(f^*)]$. The inequality in (5.2) holds because $U_c \geq U$, and the “if and only if” part is clear from the definition of $\{U < U_c\}$.

- (4) Finally, to show (5.6), recall that $P[S_{i2}] = E[(1 - \lambda)1_{S_i}]$. Subtracting $E[(1 - \lambda)1_{S_{i2}}]$ on both sides gives

$$E[\lambda 1_{S_{i2}}] = E[(1 - \lambda)1_{S_i}] - E[(1 - \lambda)1_{S_{i2}}] = E[(1 - \lambda)1_{S_{i1}}].$$

This equality can be combined with (5.5) to deduce

$$E[\varphi \lambda 1_{S_{i2}}] \leq \sup_{\omega \in S_{i2}} \varphi(\omega) E[\lambda 1_{S_{i2}}] \leq \inf_{\omega \in S_{i1}} \varphi(\omega) E[(1 - \lambda)1_{S_{i1}}] \leq E[\varphi(1 - \lambda)1_{S_{i1}}].$$

Because $a_i \leq b_i$, the product of the positive terms $E[\varphi(1 - \lambda)1_{S_{i1}}] - E[\varphi \lambda 1_{S_{i2}}]$ and $b_i - a_i$ is again positive, and this can be rewritten as

$$a_i E[\varphi(1 - \lambda)1_{S_{i1}}] + b_i E[\varphi \lambda 1_{S_{i2}}] \leq a_i E[\varphi \lambda 1_{S_{i2}}] + b_i E[\varphi(1 - \lambda)1_{S_{i1}}]. \quad (5.7)$$

Finally, rewriting $a_i E[\varphi 1_{S_{i1}}] + b_i E[\varphi 1_{S_{i2}}]$ in terms of λ , applying (5.7) and rewriting the resulting terms again in a compact form yields

$$\begin{aligned} & a_i E[\varphi 1_{S_{i1}}] + b_i E[\varphi 1_{S_{i2}}] \\ &= a_i E[\varphi \lambda 1_{S_{i1}}] + a_i E[\varphi(1 - \lambda)1_{S_{i1}}] + b_i E[\varphi \lambda 1_{S_{i2}}] + b_i E[\varphi(1 - \lambda)1_{S_{i2}}] \\ &\leq a_i E[\varphi \lambda 1_{S_{i1}}] + a_i E[\varphi \lambda 1_{S_{i2}}] + b_i E[\varphi(1 - \lambda)1_{S_{i1}}] + b_i E[\varphi(1 - \lambda)1_{S_{i2}}] \\ &= E[\varphi(\lambda a_i + (1 - \lambda)b_i)1_{S_i}], \end{aligned}$$

which finishes the proof of (5.6). \square

Theorem 5.1 follows now directly from Proposition 5.3, as follows.

Proof of Theorem 5.1 The inequality “ \leq ” for (5.1) follows from $U \leq U_c$. For “ \geq ”, we start with some $f \in C(x)$. Proposition (5.3) gives $f^* \in C(x)$ with $E[U(f^*)] = E[U_c(f)]$. The inequality “ \geq ” follows since $f \in C(x)$ is arbitrary.

For the second part, we first assume that there is a maximizer $f \in C(x)$ for the problem $u(x, U)$. Then (5.1) gives

$$u(x, U) = E[U(f)] \leq E[U_c(f)] \leq u(x, U_c) = u(x, U),$$

which shows that f also maximizes $u(x, U_c)$. Conversely, assume that f maximizes $u(x, U_c)$. Proposition 5.3 gives $f^* \in C(x)$ satisfying $E[U(f^*)] = E[U_c(f^*)] = E[U_c(f)]$. We deduce that

$$u(x, U) = u(x, U_c) = E[U_c(f)] = E[U_c(f^*)] = E[U(f^*)],$$

which shows that f^* is a maximizer for $u(x, U)$. \square

Remark 5.4 Looking at the proof of Proposition 5.3 more closely shows that we have not directly used the fact that the U is defined on \mathbb{R}_+ . The essential ingredients are that U_c exists and that $\{U < U_c\}$ can be written as a countable (finite) union of bounded intervals. The proof is also valid for utilities defined on \mathbb{R} if these two assumptions are satisfied.

The first message of Theorem 5.1 is that, in the atomless case, any solution for the optimization problem $u(x, U)$ also solves the concavified one, which in turn is well understood (see, for instance, Bouchard et al. [14] and Westray and Zheng [41]). In particular, the structure of the solution for the concavified problem is known. In this way, Theorem 5.1 can be used to describe explicitly the structure of all solutions to the optimization problem $u(x, U)$. Theorem 5.1 also says that the existence of a maximizer for the concavified optimization problem already guarantees the existence of a maximizer for the problem $u(x, U)$. Since

the conditions for existence for the concavified problem are slightly weaker than those in Theorems 3.3 and 5.1 can be used to slightly relax the assumptions in the present atomless case, as follows.

Corollary 5.5 *Suppose that (Ω, \mathcal{F}, P) is atomless and that $u(x, U) < \infty$ for some $x > 0$. Then there exists $\tilde{x} \in (0, \infty]$ such that for $x \in (0, \tilde{x})$, the problem $u(x, U)$ has a maximizer and every maximizer satisfies $f^* \in -\partial J(\lambda^* \varphi)$ for some $\lambda^* > 0$. If Assumption 3.1 is satisfied, then $\tilde{x} = \infty$.*

Proof The argument is a combination of known results on non-smooth utility maximization, non-smooth versions of results of Jin et al. [28], and Theorem 5.1. For completeness, we summarize the material on non-smooth utility maximization in Appendix B.

Proposition B.1 shows that there is $\tilde{x} \in (0, \infty]$ such that for $x \in (0, \tilde{x})$, the concavified optimization problem $u(x, U_c)$ has a maximizer. By Theorem 5.1, the problem $u(x, U)$ then also admits a solution f^* which also maximizes $u(x, U_c)$. The particular form of f^* then follows from Proposition B.1. If Assumption 3.1 is satisfied, the statements follow in much the same way by using Proposition B.3 instead of Proposition B.1. \square

For models with an atomless underlying probability space, Corollary 5.5 shows that any maximizer for $u(x, U)$ satisfies $f^* \in -\partial J(\lambda \varphi)$ for some $\lambda \geq 0$. But J is a convex function and the subdifferential is monotone in the sense that $J'_+(z_1) \leq J'_-(x) \leq J'_+(x) \leq J'_-(z_2)$ when $z_1 < x < z_2$ (see part (iv) of Lemma A.1). This implies that f^* and $\lambda \varphi$ are anti-comonotonic. If U_c is strictly increasing, the first order condition implies $\lambda > 0$ and we obtain:

Corollary 5.6 *Let U_c be strictly increasing, let (Ω, \mathcal{F}, P) be atomless and let f^* be a maximizer for $u(x, U)$ derived in Corollary 5.5. Then f^* and φ are anti-comonotonic.*

In some examples, the distribution of the pricing density φ is continuous. This simplifies several arguments and allows for further results. Larsen [31], Carassus and Pham [15] and Rieger [35] consider a framework with $\{U < U_c\} = \bigcup_{i=1}^N (a_i, b_i)$ for a fixed N and argue that a maximizer for $u(x, U_c)$ also maximizes $u(x, U)$. The arguments of Larsen [31] can easily be adapted to our slightly more general case.

Lemma 5.7 *Let φ have a continuous distribution and let $f^* \in -\partial J(\lambda \varphi)$ be a maximizer f^* for $u(x, U_c)$. Then f^* is also a maximizer for $u(x, U)$.*

Proof Let f^* be a maximizer for the concavified problem $u(x, U_c)$. If

$$P[f^* \in \{U < U_c\}] = 0, \quad (5.8)$$

it follows that $u(x, U) \leq u(x, U_c) = E[U_c(f^*)] = E[U(f^*)] \leq u(x, U)$ which means that f^* is also a maximizer for $u(x, U)$. To prove (5.8), note that $f^* \in -\partial J(\lambda \varphi)$ for some $\lambda \geq 0$. The case $\lambda = 0$ gives $f^* = -J'_+(0)$ and the result follows from Lemma A.2. So consider $\lambda > 0$. It is shown in Lemma 2.8 that U_c is locally affine on $\{U < U_c\}$. So if $f^*(\omega) \in \{U < U_c\}$, then $\partial U_c(f^*(\omega)) = \{\lambda \varphi(\omega)\}$ by Lemma 2.9. But if U_c is affine on an interval (a, b) with slope c , then it follows that J is not differentiable in c since $x \in -\partial J(c)$ for $x \in (a, b)$. This gives

$$\{f^* \in \{U < U_c\}\} \subseteq \{\omega \in \Omega \mid J \text{ is not differentiable in } \lambda \varphi(\omega)\}.$$

The function J is finite and convex on $(0, \infty)$ and thus differentiable there except for at most countably many points $y_i, i \in \mathbb{N}$. This yields (5.8) since

$$P[\{\omega \in \Omega \mid J(\cdot) \text{ is not differentiable in } \lambda\varphi(\omega)\}] = \sum_i P[\{\lambda\varphi = y_i\}] = 0,$$

where continuity of the distribution of φ is used in the last equality. \square

Remark 5.8 (1) If the distribution of φ is continuous, Theorem 5.1 follows directly from Lemma 5.7.

(2) Lemma 5.7 is also useful for the more general case with more than one pricing measure. In *concave* utility maximization, one central property is that the solution to the problem with more than one pricing measure and the utility U_c can be obtained by working with only one well-chosen (generalized) pricing measure (see, for instance, Kramkov and Schachermayer [29] for details). If the distribution of the (generalized) pricing density is continuous, it follows as in Lemma 5.7 that the solution f^* to the problem with this pricing measure satisfies $P[f^* \in \{U < U_c\}] = 0$. This implies that it is also the solution to the problem with multiple pricing measures and the utility U . This idea is used in He and Zhou [22] for one particular model. Bichuch and Sturm [13] explore the idea more thoroughly and derive conditions under which it can be applied for specific classes of models.

(3) The case with (infinitely) many pricing measures having not necessarily continuous distributions is more subtle. As in part 2), the solution f^* for U_c satisfies $f^* \in -\partial J(\lambda^*\varphi)$ and $E[\varphi f^*] = x$ for a particular dual object φ . However, Westray and Zheng [42] show that these conditions are in general not sufficient for optimality of f^* in the concavified problem. This is in contrast to the case with one pricing measure where these assumptions are sufficient for optimality and this is the reason why the proof of Proposition 5.3 cannot be extended directly to the case with many pricing measures.

In a similar way as in Lemma 5.7, we also get uniqueness for the maximizer for $u(x, U)$ provided that U_c is strictly increasing.

Lemma 5.9 *Let U_c be strictly increasing, let φ have a continuous distribution and let f^* be a maximizer for $u(x, U)$ derived in Corollary 5.5. Then the maximizer is P -a.s. unique.*

Proof Let $f_i \in -\partial J(\lambda_i\varphi)$ for $i = 1, 2$ be maximizers for $u(x, U)$. If $\lambda_1 > \lambda_2$, it follows from part (ix) of Lemma A.1 that $f_1 \leq f_2$. Since $f_i \in \{U = U_c\}$ P -a.s. and U is strictly increasing there, we deduce that $f_1 = f_2$ P -a.s. Suppose now $\lambda_1 = \lambda_2$. If J is differentiable in $\lambda_1\varphi(\omega)$, then $\partial J(\lambda_1\varphi(\omega))$ is a singleton and hence $f_1(\omega) = f_2(\omega)$. But this implies

$$\{f_1 \neq f_2\} \subseteq \{\omega \in \Omega \mid J(\cdot) \text{ is not differentiable in } \lambda^*\varphi(\omega)\},$$

and the latter has probability 0 since the distribution of φ is continuous. \square

We finally discuss one example which makes the advantages of Theorem 5.1 transparent.

Example 5.10 We consider some time horizon $T \in (0, \infty)$ and a probability space (Ω, \mathcal{F}, P) on which there are a Poisson process $(N_t)_{t \geq 0}$ with intensity $\lambda > 0$ and a Brownian motion $(W_t)_{t \geq 0}$. Let $(\mathcal{F}_t)_{0 \leq t \leq T}$ be the augmented filtration generated by N and W . We consider a (discounted) market consisting of a savings account $B \equiv 1$ and two risky stocks S^1, S^2 described by

$$\begin{aligned} dS_t^1 &= \alpha S_t^1 dt + \sigma_1 S_t^1 d\tilde{N}_t, & S_0^1 &= 1, \sigma_1 > -1, \sigma_1 \neq 0, \alpha/\sigma_1 < \lambda, \\ dS_t^2 &= \sigma_2 S_t^2 dW_t, & S_0^2 &= 1, \sigma_2 > 0, \end{aligned}$$

where $\tilde{N}_t := N_t - \lambda t$ is the compensated Poisson process. This model defines a complete financial market and the unique martingale measure is defined by $dQ/dP = e^{(\lambda - \tilde{\lambda})T} \left(\frac{\tilde{\lambda}}{\lambda}\right)^{N_T}$

for $\tilde{\lambda} := \lambda - \alpha/\sigma$ (for details, see for instance Sect. 10.6.1 in Jeanblanc et al. [26]). So instead of maximizing expected utility over dynamic strategies in the market, we may as well solve (2.2) for $\varphi = dQ/dP$. The underlying probability space is atomless, so it follows from Theorem 5.1 that $u(x, U) = u(x, U_c)$. However, the distribution of φ is not continuous, so $u(x, U)$ cannot be solved directly via Lemma 5.7.

Note that the martingale measure here is the same as in Example 2.5 even though the financial markets are not identical. Looking at the static problem (2.2), the difference is reflected in the underlying probability space. Here it is atomless while it is not atomless in Example 2.5; the set $\{N_T = 0\}$ is an atom in that case. The economic interpretation is as follows. The market from Example 2.5 consisting only of B and S^1 defines a complete market in continuous time in which one *cannot* generate an arbitrary probability distribution. An agent having a utility U can therefore not always generate the same expected utility as an agent with the concave utility U_c . Introducing as above the third asset S^2 does not increase the optimal expected utility for the agent with the utility U_c since dQ/dP does not change; his optimal final position is therefore the same as before. However, the new asset increases the optimal expected utility for the agent with the utility U . Because the underlying space is now atomless and the financial market is complete, the agent is able to generate by self-financing trading any distribution; in this way the agent generates the same optimal expected utility as the agent with the concave utility U_c .

To put the present section into perspective, let us finally discuss some results in the literature which belong to the class of models analyzed here. Berkelaar et al. [8], Larsen [31], Carassus and Pham [15], Rieger [35] and Basak and Makarov [5] consider utility maximization problems for not necessarily concave utility functions under the assumption that φ has a continuous distribution and use (sometimes implicitly) the argument given in Lemma 5.7. In addition, Carassus and Pham [15] consider Example 2.4 for $\mu = 0$ and use dynamic programming tools to show that a particular optimization problem satisfies $u(x, U) = u(x, U_c)$. Our Theorem 5.1 and Corollary 5.5 provide a unified derivation and perspective for these results. Moreover, our result also allows us to deal with other examples, such as Example 5.10, that are not possible to solve with the existing results in the literature.

6 Conclusion

In this paper, we study the problem of utility maximization from terminal wealth for a not necessarily concave utility function and for a budget set given by a single pricing measure. This is motivated by several applications in finance such as manager compensation, portfolio delegation and behavioural finance. In the literature, the classical approach is to reduce the problem to a concavified problem and to apply the classical techniques from the concave case. While this approach is powerful for specific examples, it cannot be applied in the general case.

The present paper analyzes the utility maximization problem directly. We first show the existence and several properties of a maximizer of $u(x, U)$. As in the classical concave case, this requires some additional assumptions that can be formulated in terms of the convex conjugate J of the utility function U . In contrast to the literature, we also study the value function $u(x, U)$ in detail. In particular, we show that $u(x, U_c)$ is the concave envelope of $u(x, U)$, which gives the relation between the maximization problem $u(x, U)$ and the concavified problem $u(x, U_c)$.

In contrast to the concave case, the underlying probability space crucially affects the optimization problem. While the value function $u(x, U)$ is not necessarily concave in general, it is shown to be concave and to equal $u(x, U_c)$ if the underlying probability space is atomless. This also gives a structural, easily verifiable assumption under which one can use the concavified problem $u(x, U_c)$ to analyze $u(x, U)$.

The significant impact of the underlying probability space raises a natural *stability* question: Do small deviations from the atomless structure change the results in the atomless model drastically? As an example, take a sequence of binomial models approximating the Black–Scholes model and for every model, consider the utility maximization problem for a fixed (not necessarily concave) utility U . It is then shown in Reichlin [33] that the sequences of value functions and optimal final positions converge to the corresponding quantities in the limit model. These results complement the present paper and provide additional intuition for the optimal behaviour of agents with a not necessarily concave utility function.

Finally, we have to point out that our results depend crucially on the setup with one single pricing measure. So our analysis can be seen as a first step which provides some fundamental intuition about agents with not necessarily concave utility functions. It remains an interesting and challenging problem to study the case with infinitely many pricing measures.

Acknowledgments I am grateful to Stephan Sturm, Jerome Detemple, Kasper Larsen, Marcel Nutz, Marc Rieger, Martin Schweizer and an anonymous referee for careful reading and helpful suggestions. Financial support by the National Centre of Competence in Research “Financial Valuation and Risk Management” (NCCR FINRISK) is gratefully acknowledged.

Appendix

A. Facts from convex analysis

Let $f : \mathbb{R} \rightarrow \bar{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ be an extended real-valued function. The *effective domain* of f , denoted by $\text{dom}(f)$, consists of all $x \in \mathbb{R}$ such that $f(x) < \infty$, and its interior is denoted by $\text{int}(\text{dom}(f))$. The function f is called *proper* if both $\text{dom}(f) \neq \emptyset$ and $f(x) > -\infty$ for all x . The *conjugate* of f is the extended real-valued function f^* on \mathbb{R} defined by

$$f^*(y) := \sup_{x \in \mathbb{R}} (xy - f(x))$$

for all $y \in \mathbb{R}$. The *biconjugate* f^{**} is defined by $f^{**} := (f^*)^*$. If f is proper, lower-semicontinuous and convex, then its *subdifferential* ∂f is the multivalued mapping defined by $\partial f(x) = \emptyset$ if $f(x) = \infty$ and

$$\partial f(x) := \{y \in \mathbb{R} \mid f(\xi) \geq f(x) + y(\xi - x) \text{ for all } \xi \in \mathbb{R}\} \quad \text{if } x \in \text{dom}(f).$$

The *convex envelope* \bar{f} of f is the largest convex function $\bar{f} \leq f$.

Lemma A.1 *Suppose that f is proper and lower-semicontinuous and its convex envelope \bar{f} is proper as well. Then:*

- (i) f^* is convex, proper and lower-semicontinuous.
- (ii) f and its convex envelope have the same conjugate.
- (iii) f^{**} is the lower-semicontinuous envelope of the convex envelope \bar{f} .
- (iv) Fix $x_0 \in \text{dom}(f)$ and $y_0 \in \partial f^{**}(x_0)$. If the conjugate f^* is differentiable in y_0 , then $f(x_0) = f^{**}(x_0)$.

Let f , in addition, be convex and extend the right and left derivative functions f'_+ and f'_- beyond the interval $\text{dom}(f)$ by setting both $= \infty$ for points lying to the right of $\text{dom}(f)$ and both $= -\infty$ for points lying to the left. Then:

- (v) $y \in \partial f(x)$ if and only if $x \in \partial f^*(y)$.
- (vi) $y \in \partial f(x)$ if and only if $f(x) + f^*(y) = xy$.
- (vii) Let $x \in \text{dom}(f)$. f has a unique subgradient at x if and only if f is differentiable at x .
- (viii) $\partial f(x) = \{y \in \mathbb{R} \mid f'_-(x) \leq y \leq f'_+(x)\}$.
- (ix) $f'_+(z_1) \leq f'_-(x) \leq f'_+(x) \leq f'_-(z_2)$ when $z_1 < x < z_2$.

Proof Most properties can be found in Rockafellar [36]; see Theorem 23.5 for (v) and (vi) Theorem 25.1 for (vii) and Theorem 24.1 and its discussion for (viii) and (ix). Statements (i), (ii) and (iii) are part of Theorem 11.1 of Rockafellar and Wets [37]. Part (iv) can be found in a similar form in Theorem 2 of Strömberg [39]. \square

We now apply Lemma A.1 with $f(x) := -U(-x)$ to prove Lemmas 2.8, 2.9 and 2.10 and derive some additional properties. For convenience, we write $\text{dom}(f)$, $\text{dom}(f^{**})$ and ∂f^{**} in terms of the concave envelope U_c , i.e., $\text{dom}(U) := \{x \in \mathbb{R} \mid U(x) > -\infty\}$, $\text{dom}(U_c) := \{x \in \mathbb{R} \mid U_c(x) > -\infty\}$ and

$$\partial U_c(x) := \{p \in \mathbb{R} \mid U_c(z) \leq U_c(x) + p(z - x) \text{ for all } z \in \mathbb{R}\}$$

for $x \in \text{dom}(U_c)$. The functions $-U(-x)$ and $-U_c(-x)$ are both lower-semicontinuous on \mathbb{R} and proper. It therefore follows from part (iii) of Lemma A.1 that $-U_c(-x) = f^{**}(x)$. Finally, note that $J(y) = f^*(y)$; so the convexity of J follows from part (i) of Lemma A.1.

Proof of Lemma 2.8 Basic properties of U_c : The condition (2.1) and the assumption $U(\infty) > 0$ imply that for any $\epsilon > 0$, there is $x_0 > 0$ such that $U(x) < \epsilon x$ for all $x \geq x_0$ and $U(x_0) > 0$. Therefore, $g(x) := U(x_0) + \epsilon x$ dominates U . Clearly, g is finite and concave. Hence $U_c \leq g$ and we conclude that U_c is on $(0, \infty)$ finite and concave, hence continuous. Moreover, the same argument also gives $U_c(x)/x \leq \epsilon + U(x_0)/x$ for $x \geq x_0$ which shows that U_c satisfies condition (2.1) as well.

U_c is locally affine on $\{U < U_c\}$: If $\{U < U_c\}$ is empty, there is nothing to prove. Otherwise choose $x_0 \in \{x \in \mathbb{R}_+ \mid U(x) < U_c(x)\}$ and fix $y_0 \in \partial U_c(x_0)$. Let us first consider the case that $y_0 \in (0, \infty)$. By way of contradiction, assume that J is differentiable in y_0 . By Lemma A.1, (iv), it then follows that $U(x_0) = U_c(x_0)$. But x_0 is chosen in such a way that $U(x_0) < U_c(x_0)$. So J is not differentiable in y_0 and it follows from Lemma A.1, (vii), that $\partial J(y_0)$ is multivalued. Thus, we have $y_0 \in \partial U_c(x)$ for all $x \in (-J'_+(y_0), -J'_-(y_0))$, which means that U_c is affine on that interval which also contains x_0 . The case $y_0 \in \text{dom}(J) \setminus (0, \infty)$ corresponds to $y_0 = 0 \in \partial U_c(x_0)$, which means that U_c is constant (hence affine) on $(-J'_+(0), \infty)$ and $-J'_+(0) \leq x_0$.

Structure of $\{U < U_c\}$: The set $\{U < U_c\} = \{x \in \mathbb{R}_+ \mid U(x) - U_c(x) < 0\}$ is open since $U - U_c$ is upper-semicontinuous. It can therefore be written as a union of disjoint open intervals. By way of contradiction, assume that $\{U < U_c\}$ contains (x_0, ∞) for some $x_0 \geq 0$. Then $U_c(x_0) = U(x_0)$ and U_c is affine on (x_0, ∞) . This contradicts the growth condition (2.1). \square

Proof of Lemma 2.9 Monotonicity of J as well as $J > -\infty$ on $(0, \infty)$ and $J = \infty$ on $(-\infty, 0)$ follow from the definition. The fact that $J < \infty$ on $(0, \infty)$ follows from the growth condition (2.1). The equivalence follows from Lemma A.1, (v) and (vi). Lemma A.1 (ii) gives that U and U_c have the same convex conjugate. \square

Lemma A.2 $U(-J'_\pm(y)) = U_c(-J'_\pm(y))$ for $y > 0$. If $-J'_+(0) < \infty$, then also $U(-J'_+(0)) = U_c(-J'_+(0))$.

Proof Recall that J is decreasing on $(0, \infty) = \text{int}(\text{dom}(J))$. So $-J'_\pm(y)$ is positive and finite. By way of contradiction, assume that there is $y > 0$ with $U(-J'_\pm(y)) < U_c(-J'_\pm(y))$. By Lemma 2.8, $\{U < U_c\} = \bigcup_i (a_i, b_i)$ is open. Hence there are i, z_1, z_2 such that $a_i < z_1 < -J'_\pm(y) < z_2 < b_i$. By Lemma 2.8, U_c is affine on (a_i, b_i) and by Lemma 2.9, we know that $U'_c(-J'_\pm(y)) = y$. This means that $z_i \in -\partial J(y)$ which contradicts Lemma A.1, (viii). If $-J'_+(0) < \infty$, the same idea gives $U(-J'_+(0)) = U_c(-J'_+(0))$. \square

Lemma A.3 The convex conjugate J satisfies $\lim_{y \rightarrow \infty} \sup_{x \in -\partial J(y)} x = 0$.

Proof Since $x \leq -J'_-(y)$ for all $x \in -\partial J(y)$, it is sufficient to show that

$$\lim_{y \rightarrow \infty} -J'_-(y) = 0. \quad (\text{A.1})$$

If $\tilde{y} := \lim_{x \rightarrow 0} \sup_{y \in \partial U_c(x)} y < \infty$, we have that $\partial J(y) = 0$ for $y > \tilde{y}$. If $\lim_{x \rightarrow 0} \sup_{y \in \partial U_c(x)} y = \infty$, fix $\epsilon > 0$ and $y_0 := \sup_{y \in \partial U_c(\epsilon)} y$. For $q \in \partial U_c(\epsilon)$ and $y > y_0$, we have $q \leq y_0 < y$. Lemma 2.9 gives $-\epsilon \in \partial J(q)$ and Lemma A.1, (ix) implies $-J'_-(y) \leq -J'_-(y_0) \leq \epsilon$, which finishes the proof of (A.1). \square

It remains to prove Lemma 2.10. The equivalence is proved in Lemma 4.1 of Deelstra et al. [19]. We only prove the second part.

Proof of Lemma 2.10 We argue that there are $x_0 > 0$ and $\gamma < 1$ such that

$$0 \leq U_c(x) \leq U(x) + \gamma U_c(x) \quad \text{on } (x_0, \infty). \quad (\text{A.2})$$

The result then follows for $k = 1/(1 - \gamma)$.

Proposition 4.1. of Deelstra et al. [19] shows that $\text{AE}_0(J) < \infty$ and the growth condition (2.1) imply the existence of two constants x_0 and $\gamma < 1$ such that $\sup_{y \in \partial U_c(x)} yx \leq \gamma U_c(x)$ holds on (x_0, ∞) . By moving x_0 to the right if necessary, we may assume that $U(x_0)$ is positive. Moreover, recall that U_c is locally affine on $\{U < U_c\} = \bigcup_i (a_i, b_i)$. So, for $x \in (a_i, b_i) \cap (x_0, \infty)$, we rewrite $U_c(x)$, use that $U_c(a_i) = U(a_i) \leq U(x)$ and apply for $y = U'_c(x)$ the above inequality $U'_c(x)x \leq \gamma U_c(x)$ to get

$$U_c(x) = U_c(a_i) + U'_c(x)(x - a_i) \leq U(x) + U'_c(x)x \leq U(x) + \gamma U_c(x).$$

For $x \in \{U = U_c\} \cap (x_0, \infty)$, (A.2) follows since $U(x) = U_c(x)$ is positive. \square

B. Non-smooth utility maximization

This appendix contains the results on non-smooth (concave) utility maximization which are relevant for the proofs in Sect. 5. It combines non-smooth versions of results of Jin et al. [28] and known results on non-smooth utility maximization as presented in Bouchard et al. [14] and Westray and Zheng [41]. We use the same notation as in Sects. 3 and 4.

Theorem 3.3 shows that Assumption 3.1 is sufficient to guarantee the existence of a maximizer for $u(x, U_c)$ for all $x > 0$. Under the weaker assumption that $u(x, U_c) < \infty$ for some $x > 0$, the existence of a maximizer is still obtained for some $x > 0$.

Proposition B.1 Assume that $u(x, U_c) < \infty$ for some $x > 0$. Then there exists $\tilde{x} > 0$ such that for $x \in (0, \tilde{x})$, the problem $u(x, U_c)$ admits a maximizer. Every maximizer f^* is of the form $f^* \in -\partial J(\lambda^* \varphi)$ for some $\lambda^* > 0$.

This result as well as its proof are non-smooth versions of Corollary 3.1 and its proof in Jin et al. [28]. The proof is based on the following lemma which is the non-smooth version of Theorem 3.1 of Jin et al. [28].

Lemma B.2 *If $E[-J'_+(\lambda\varphi)\varphi] = \infty$ for some $\lambda > 0$, there is some $x > 0$ such that $u(x, U_c) = \infty$.*

Proof (1) We start with the case $U_c(0) := \lim_{x \rightarrow 0} U_c(x) \geq 0$. Fix some $\lambda_0 > 0$ and $x > 0$. Since $E[-J'_+(\lambda_0\varphi)\varphi] = \infty$, one can find a set $A \in \mathcal{F}$ such that $E[-J'_+(\lambda_0\varphi)\varphi 1_A] \in (x, \infty)$. Define $h(\lambda) := E[-J'_+(\lambda\varphi)\varphi 1_A]$ for $\lambda \in [\lambda_0, \infty)$. Since $-J'_-$ is decreasing, the function $h(\lambda)$ is decreasing. Moreover, the monotone convergence theorem and Lemma A.3 yield $\lim_{\lambda \rightarrow \infty} h(\lambda) = 0$. Therefore, there is some $\lambda_1 \geq \lambda_0$ such that

$$E[-J'_+(\lambda_1\varphi)\varphi 1_A] = h(\lambda_1+) \leq x \leq h(\lambda_1-) = E[-J'_-(\lambda_1\varphi)\varphi 1_A].$$

Thus we find f^* satisfying $E[\varphi f^*] = x$ and $f^* \in -\partial J(\lambda_1\varphi)$ on A . Lemma 2.9 gives $U_c(f^*(\omega)) = J(\lambda_1\varphi(\omega)) + f^*(\omega)\lambda_1\varphi(\omega)$ for $\omega \in A$ and $U_c(f^*(\omega)) = U_c(0)$ for $\omega \in A^c$. The assumption $U_c \geq 0$ implies $J \geq 0$ and we conclude that

$$E[U_c(f^*)] \geq E[(J(\lambda_1\varphi) + f^*\lambda_1\varphi)1_A] \geq \lambda_1 E[\varphi f^* 1_A] = \lambda_1 x \geq \lambda_0 x.$$

Since $\lambda_0 > 0$ is arbitrary, we arrive at $u(x, U_c) \geq \lim_{\lambda \rightarrow \infty} \lambda x = \infty$.

(2) For the general case, define a shifted utility $U_c^{x_0}(x) := U_c(x_0 + x)$ on \mathbb{R}_+ , where $x_0 := \inf\{x > 0 \mid U(x) \geq 0\}$, and denote the convex conjugate of $U_c^{x_0}$ by J^{x_0} . If we show (as we do below) that

$$E[-J'_+(\lambda\varphi)\varphi] = \infty \quad \forall \lambda > 0 \implies E[-(J^{x_0})'_+(\lambda\varphi)\varphi] = \infty \quad \forall \lambda > 0, \quad (\text{B.1})$$

we can use the fact that $U_c^{x_0} \geq 0$ and part 1) to see that

$$u(\tilde{x}, U_c) \geq \sup\{E[U_c(f)] \mid f \in L_+^0, f \geq x_0, E[\varphi f] \leq \tilde{x}\} = \infty$$

for \tilde{x} large enough, which gives the statement of the lemma.

It remains to prove (B.1). In order to relate ∂J^{x_0} and ∂J , note that

$$J^{x_0}(y) - x_0 y = \sup_{x > 0} \{U_c(x_0 + x) - (x + x_0)y\} \leq J(y) \quad (\text{B.2})$$

holds for arbitrary $y > 0$. For $y \leq y_0 := \inf_{q \in \partial U_c(x_0)} q$, fix some $\tilde{x} \in -\partial J(y)$ and note that $\tilde{x} \geq x_0$ by Lemma A.1, (ix). Together with the conjugacy relation between \tilde{x} and y and the definition of J^{x_0} , we obtain

$$J(y) = U_c(\tilde{x}) - \tilde{x}y = U_c(x_0 + (\tilde{x} - x_0)) - (\tilde{x} - x_0)y - x_0y \leq J^{x_0}(y) - x_0y,$$

which implies that there is equality in (B.2) for $y \leq y_0$. We use this together with the definition of $(J^{x_0})'_-$ and (B.2) (in this order) to see that for $y < y_0$,

$$\begin{aligned} J(y) + ((J^{x_0})'_-(y) - x_0)(z - y) \\ = J^{x_0}(y) - x_0y + (J^{x_0})'_-(y)(z - y) - x_0(z - y) \leq J^{x_0}(z) - x_0z \leq J(z) \end{aligned}$$

holds for $z > 0$ which means that $(J^{x_0})'_-(y) - x_0 \in \partial J(y)$ for $y < y_0$. Since $E[-J'_+(\lambda\varphi)\varphi 1_{\{\lambda\varphi > y_0\}}] \leq -J'_+(y_0) < \infty$, the assumption $E[-J'_+(\lambda\varphi)\varphi] = \infty$ for all $\lambda > 0$ also implies

$$\begin{aligned} E[-J'_+(\lambda\varphi)\varphi 1_{\{\lambda\varphi \leq y_0\}}] &= E[-J'_+(\lambda\varphi)\varphi (1 - 1_{\{\lambda\varphi > y_0\}})] \\ &\geq E[-J'_+(\lambda\varphi)\varphi] - J'_+(y_0) = \infty \end{aligned} \quad (\text{B.3})$$

for all $\lambda > 0$. But then, plugging $(J^{x_0})'_-(y) - x_0 \in \partial J(y)$ into (B.3) gives

$$\begin{aligned} \infty &= E[-J'_+(\lambda\varphi)\varphi 1_{\{\lambda\varphi \leq y_0\}}] \leq E[-((J^{x_0})'_-(\lambda\varphi) - x_0)\varphi 1_{\{\lambda\varphi \leq y_0\}}] \\ &\leq E[-(J^{x_0})'_-(\lambda\varphi)\varphi] + x_0 \end{aligned}$$

for every $\lambda > 0$, which finishes the proof of (B.1). \square

Proof of Proposition B.1 The assumption $u(x, U_c) < \infty$ for some $x > 0$ and the concavity of $u(\cdot, U_c)$ imply that $u(x, U_c) < \infty$ for all $x > 0$. Thus we apply Lemma B.2 to get $\lambda_0 > 0$ satisfying $\tilde{x} := E[-J'_+(\lambda_0\varphi)\varphi] < \infty$. Similarly to the proof of Lemma B.2, define $h(\lambda) := E[-J'_+(\lambda\varphi)\varphi]$ for $\lambda \in [\lambda_0, \infty)$. Since $-J'_+$ is decreasing, the function $h(\lambda)$ is decreasing. Moreover, the monotone convergence theorem and Lemma A.3 yield $\lim_{\lambda \rightarrow \infty} h(\lambda) = 0$. Hence for all $x \in (0, \tilde{x}]$, there is some $\lambda_1 \geq \lambda_0$ such that

$$E[-J'_+(\lambda_1\varphi)\varphi] = h(\lambda_1+) \leq x \leq h(\lambda_1-) = E[-J'_-(\lambda_1\varphi)\varphi].$$

Thus we find f^* satisfying $E[\varphi f^*] = x$ and $f^* \in -\partial J(\lambda_1\varphi)$. In order to show optimality of f^* , consider $f \in C(x)$ and note that $E[\varphi f] \leq x$ by the definition of $C(x)$. It follows from the definition of J , Lemma 2.9 and $E[\varphi f^*] = x$ that

$$\begin{aligned} E[U_c(f)] &= E[U_c(f) - \lambda_1 f\varphi] + E[\lambda_1 f\varphi] \\ &\leq E[J(\lambda_1\varphi)] + \lambda_1 x \\ &= E[U_c(f^*)] - E[f^*\varphi\lambda_1] + \lambda_1 x = E[U_c(f^*)]. \end{aligned} \quad (\text{B.4})$$

Hence $u(x, U_c) = E[U_c(f^*)]$ since $f \in C(x)$ was arbitrary. Note that the inequality in (B.4) is strict if $f \notin -\partial J(\lambda_1\varphi)$ on a set with strictly positive measure. Thus any other maximizer \hat{f} satisfies $\hat{f} \in -\partial J(\lambda_1\varphi)$ as well. \square

Under Assumption 3.1, Theorem 3.2 in Bouchard et al. [14] or Theorem 5.1 in Westray and Zheng [41] give that Proposition B.1 holds for any $x > 0$:

Proposition B.3 *Suppose Assumption 3.1 is satisfied. Then the concave problem $u(x, U_c)$ has a solution $f^* \in C(x)$ for every $x > 0$. Every solution satisfies $f^* \in -\partial J(\lambda^*\varphi)$ for some $\lambda^* \geq 0$.*

C. Proof of Lemma 3.2

Proof The fact that Assumption 3.1 implies $u(x, U) < \infty$ for $x > 0$ is straightforward since we have $E[U(f)] \leq E[U_c(f)] \leq E[J(\lambda\varphi)] + \lambda x_0$ for all $f \in C(x_0)$ and $\lambda > 0$. For the converse direction, we use that $\text{AE}_0(J) < \infty$ and $u(x, U_c) < \infty$ for some $x > 0$ imply Assumption 3.1 (see for instance Lemma 5.4 of Westray and Zheng [41]). It is therefore sufficient to show that $\text{AE}_0(J) < \infty$ and $u(x, U) < \infty$ for some $x > 0$ imply $u(x, U_c) < \infty$ for some $x > 0$. For this, let $x_0 > 0$ and k be as given in the second part of Lemma 2.10 so that

$$0 \leq U_c(x) \leq kU(x) \quad \text{on } (x_0, \infty). \quad (\text{C.1})$$

If U is positive on $(0, \infty)$, then fix some $f \in C(x)$ and apply (C.1) on the set $\{f > x_0\}$. This gives $E[U_c(f)] \leq U_c(x_0) + kE[U(f)]$ and taking the supremum over all $f \in C(x)$ implies

$u(x, U_c) \leq U_c(x_0) + ku(x, U)$. So if $u(x, U)$ is finite for some $x > 0$, then $u(x, U_c)$ is also finite.

If $U(0) < 0$, we choose ϵ small enough such that $x - \epsilon > 0$, fix $f \in C(x - \epsilon)$ and apply the above argument to $f_\epsilon := f + \epsilon$ and $U(x) - U(\epsilon)$. This gives

$$\begin{aligned} E[U_c(f)] &\leq E[U_c(f_\epsilon)] \\ &\leq U_c(x_0) + kE[(U(f_\epsilon) - U(\epsilon))1_{\{f_\epsilon \geq x_0\}}] + kE[U(\epsilon)1_{\{f_\epsilon \geq x_0\}}] \\ &\leq U_c(x_0) + kE[U(f_\epsilon) - U(\epsilon)] + kE[U(\epsilon)1_{\{f_\epsilon \geq x_0\}}] \\ &\leq U_c(x_0) + ku(x, U) + k|U(\epsilon)|, \end{aligned}$$

where $f_\epsilon \in C(x)$ is used in the last step. Taking the supremum over all $f \in C(x - \epsilon)$ gives then $u(x - \epsilon, U_c) \leq ku(x, U) + \tilde{k}$ for some constant \tilde{k} and the result follows. \square

References

- Aliprantis, C.D., Border, K.C.: Infinite Dimensional Analysis: A Hitchhiker's Guide, 3rd edn. Springer, Berlin (2006)
- Aumann, R.J., Perles, M.: A variational problem arising in economics. *J. Math. Anal. Appl.* **11**, 488–503 (1965)
- Balder, E.J.: A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM J. Control Optim.* **22**, 570–598 (1984)
- Balder, E.J., Pistorius, M.R.: On an optimal consumption problem for p -integrable consumption plans. *Econ. Theory* **17**, 721–737 (2001)
- Basak, S., Makarov, D.: Strategic asset allocation in money management. *J. Finance* (2011). Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1336546
- Beare, B.K.: Measure preserving derivatives and the pricing kernel puzzle. *J. Math. Econ.* **47**, 689–697 (2011)
- Benartzi, S., Thaler, R.H.: Myopic loss aversion and the equity premium puzzle. *Q. J. Econ.* **110**, 73–92 (1995)
- Berkelaar, A.B., Kouwenberg, R., Post, T.: Optimal portfolio choice under loss aversion. *Rev. Econ. Stat.* **86**, 973–987 (2004)
- Bernard, C., Boyle, P.P.: Explicit representation of cost-efficient strategies. University of Waterloo (2010). Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1695543
- Bernard, C., Ghossoub, M.: Static portfolio choice under cumulative prospect theory. *Math. Financial Econ.* **2**, 277–306 (2010)
- Biagini, S.: Expected utility maximization: Duality methods. In: Cont, R. (ed.) *Encyclopedia of Quantitative Finance*, pp. 638–645. Wiley, Chichester (2010)
- Biagini, S., Guasoni, P.: Relaxed utility maximization in complete markets. *Math. Finance* **21**, 703–722 (2011)
- Bichuch, M., Sturm, S.: Portfolio optimization under convex incentive schemes. Princeton University (2011). Available at <http://arxiv.org/abs/1109.2945>
- Bouchard, B., Touzi, N., Zeghal, A.: Dual formulation of the utility maximization problem: the case of nonsmooth utility. *Ann. Appl. Probab.* **14**, 678–717 (2004)
- Carassus, L., Pham, H.: Portfolio optimization for piecewise concave criteria. In: Ogawa, S. (ed.) *The 8th Workshop on Stochastic Numerics*, volume 1620 of RIMS Kôkyûroku series, pp. 81–108. Research Institute for Mathematical Sciences, Kyoto (2009)
- Carassus, L., Rásonyi, M.: On optimal investment for a behavioural investor in multiperiod incomplete market models. University of Edinburgh (2011). Available at <http://arxiv.org/abs/1107.1617>
- Carlier, G., Dana, R.A.: Optimal demand for contingent claims when agents have law invariant utilities. *Math. Finance* **21**, 169–201 (2011)
- Carpenter, J.N.: Does option compensation increase managerial risk appetite? *J. Finance* **55**, 2311–2331 (2000)
- Deelstra, G., Pham, H., Touzi, N.: Dual formulation of the utility maximization problem under transaction costs. *Ann. Appl. Probab.* **11**, 1353–1383 (2001)
- Dybvig, P.H.: Distributional analysis of portfolio choice. *J. Bus.* **61**, 369–393 (1988)

21. Föllmer, H., Schied, A.: *Stochastic Finance: An Introduction in Discrete Time*, 3rd edn. Walter de Gruyter, Berlin (2011)
22. He, X.D., Zhou, X.Y.: Portfolio choice via quantiles. *Math. Finance* **21**, 203–231 (2011)
23. He, X.D., Zhou, X.Y.: Portfolio choice under cumulative prospect theory: an analytical treatment. *Manag. Sci.* **57**, 315–331 (2011)
24. Hens, T., Reichlin, C.: Three solutions to the pricing kernel puzzle. *Rev. Finance* (2012). Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1582888
25. Jackwerth, J.C.: Recovering risk aversion from option prices and realized returns. *Rev. Financial Stud.* **13**, 433–451 (2000)
26. Jeanblanc, M., Yor, M., Chesney, M.: *Mathematical Methods for Financial Markets*. Springer, Berlin (2009)
27. Jin, H., Zhou, X.Y.: Behavioral portfolio selection in continuous time. *Math. Finance* **18**, 385–426 (2008)
28. Jin, H., Xu, Z.Q., Zhou, X.Y.: A convex stochastic optimization problem arising from portfolio selection. *Math. Finance* **18**, 171–183 (2008)
29. Kramkov, D., Schachermayer, W.: The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Ann. Appl. Probab.* **9**, 904–950 (1999)
30. Kulldorff, M.: Optimal control of favorable games with a time limit. *SIAM J. Control Optim.* **31**, 52–69 (1993)
31. Larsen, K.: Optimal portfolio delegation when parties have different coefficients of risk aversion. *Quant. Finance* **5**, 503–512 (2005)
32. Pliska, S.R.: A stochastic calculus model of continuous trading: optimal portfolios. *Math. Oper. Res.* **11**, 371–382 (1986)
33. Reichlin, C.: Behavioural portfolio selection: asymptotics and stability along a sequence of models. ETH, Zurich (2012)
34. Rieger, M.O.: Co-monotonicity of optimal investments and the design of structured financial products. *Finance Stoch.* **15**, 27–55 (2011)
35. Rieger, M.O.: Optimal financial investments for non-concave utility functions. *Econ. Lett.* **114**, 239–240 (2012)
36. Rockafellar, R.T.: *Convex Anal.* Princeton University Press, Princeton, NJ (1972)
37. Rockafellar, R.T., Wets, R.J-B.: *Variational Analysis*. Springer, Berlin (1998)
38. Siorpaes, P.: Do arbitrage-free prices come from utility maximization? University of Vienna (2012). Available at <http://arxiv.org/pdf/1207.4749.pdf>
39. Strömberg, T.: A note on the differentiability of conjugate functions. *Archiv der Mathematik* **93**, 481–485 (2009)
40. Tversky, A., Kahneman, D.: Advances in prospect theory: cumulative representation of uncertainty. *J. Risk Uncertain.* **5**, 297–323 (1992)
41. Westray, N., Zheng, H.: Constrained nonsmooth utility maximization without quadratic inf convolution. *Stoch. Process. Appl.* **119**, 1561–1579 (2009)
42. Westray, N., Zheng, H.: Minimal sufficient conditions for a primal optimizer in nonsmooth utility maximization. *Finance Stoch.* **15**, 501–512 (2011)